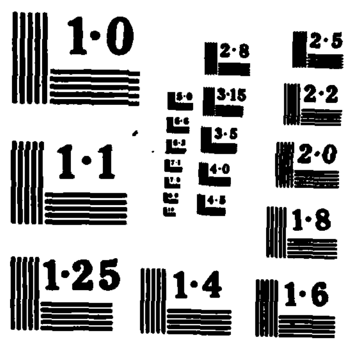
AD-A154 302 COMMENTS ON ESTIMATING QUANTILES FOR GAUSSIAN FUNCTIONALS BY SIMULATION(U) VIRGINIA UNIV CHARLOTTESVILLE DEPT OF SYSTEMS ENGINEERING C M HARRIS 1/1

UNCLASSIFIED APR 85 UVA/525393/SE85/109 N00014-83-K-0624 F/G 12/1 NL

END

FILMED

DTIC


1·0
2·8
2·5
3·15
2·2
3·5
2·0
1·1
4·0
4·5
1·8
1·25
1·4
1·6

AD-A154 302

A Technical Report

Grant No. N00014-83-K-0624

COMMENTS ON ESTIMATING QUANTILES FOR GAUSSIAN FUNCTIONALS BY SIMULATION

Submitted to:

Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Attention: Program Manager,
Statistics and Probability

Submitted by:

C. M. Harris
Principal Investigator

Serial NH-1

Contract N00014-83-K-0624
Task B; Project NR 347-139

Report No. UVA/525393/SE85/109
April 1985

SCHOOL OF ENGINEERING AND APPLIED SCIENCE

DEPARTMENT OF SYSTEMS ENGINEERING

UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22901

A Technical Report

Grant No. N00014-83-K-0624

COMMENTS ON ESTIMATING QUANTILES FOR GAUSSIAN FUNCTIONALS BY SIMULATION

Submitted to:

Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Attention: Program Manager,
Statistics and Probability

Submitted by:

C. M. Harris
Principal Investigator

Serial NH-1

Contract N00014-83-K-0624
Task B; Project NR 347-139

Department of Systems Engineering
SCHOOL OF ENGINEERING AND APPLIED SCIENCE
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA

Report No. UVA/525393/SE85/109
April 1985

Copy No. _____

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. A154 302
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) Comments on Estimating Quantiles for Gaussian Functionals by Simulation	5. TYPE OF REPORT & PERIOD COVERED Technical Report
	6. PERFORMING ORG. REPORT NUMBER UVA/525393/SE85/109
7. AUTHOR(s) Carl M. Harris	8. CONTRACT OR GRANT NUMBER(s) N00014-83-K-0624
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Systems Engineering, University of Virginia, Thornton Hall, Charlottesville, VA 22901	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Task B; Project NR347-139
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, 800 North Quincy Street, Arlington, VA 22217	12. REPORT DATE April 1985
	13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE Unclassified

16. DISTRIBUTION STATEMENT (of this Report)

Unlimited

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

simulation; quantile estimation; distribution function estimation; Gaussian functionals; Brownian motion.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The construction of an optimum model-sampling simulation experimental design is often an art and clearly very much problem dependent. The accurate estimation of the distribution of any random variable by computer is almost always difficult, especially when reasonable accuracy is desired for its tail probabilities. If, in addition, each element of the computer generated pseudo-random sequence is itself the result of a stochastic limit or possibly a functional of a continuous-time process, it becomes most complicated to assess the final statistics. In this paper, we focus on the estimation of tail

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (continued)

probabilities for the distribution function of the maximum on the unit interval of a continuous-time Wiener process approximated as a multivariate normal of increasing dimension. We critique recent approaches to sample-size determination for such distribution-sampling problems and build on insights from probability theory to find more reasonable run sizes.

Abstract

The construction of an optimum model-sampling simulation experimental design is often an art and clearly very much problem dependent. The accurate estimation of the distribution of any random variable by computer is almost always difficult, especially when reasonable accuracy is desired for its tail probabilities. If, in addition, each element of the computer-generated pseudo-random sequence is itself the result of a stochastic limit or possibly a functional of a continuous-time process, it becomes most complicated to assess the final statistics. In this paper, we focus on the estimation of tail probabilities for the distribution function of the maximum on the unit interval of a continuous-time Wiener process approximated as a multivariate normal of increasing dimension. We critique recent approaches to sample-size determination for such distribution-sampling problems and build on insights from probability theory to find more reasonable run sizes.

Keywords include:
-> see 1473

1. Introduction

The determination of a best run size for a model-sampling simulator is generally a nontrivial matter. Many authors (for example, see Bratley, Fox and Schrage, 1983, and Rubinstein, 1981) have commented on the difficulty of accurately estimating measures of random variables using computers even when generating random (or pseudo-random) samples. Limit theorems are typically

of little use when the underlying probabilistic structure is complex and rates of convergence not available. Furthermore, it is often quite hard to obtain appropriate estimates of variability for calculating confidence intervals or making statements of accuracy.

It is well known that run lengths necessary to achieve desired levels of "precision" for estimating such things as population moments and quantiles may rise very quickly with the order of the percentile or moment. With the advent of inexpensive and accessible computing, resource and time constraints are generally not at issue (though repeated sorting may be necessary). Instead, it often becomes more important to understand the behavior of the population parameters sought. This sort of problem becomes even more complicated, when the values of each element of an IID sequence are themselves the result of a further limiting process, (as, for example, in extreme-value problems), so that the full sequence is actually one small series imbedded within another larger one.

Excellent examples of these problems arise in the use of simulation methods for estimating functionals of Gaussian processes. We draw attention to the work of Serfling and Wood (1976), Wood (1978), Siegmund (1978), and Chandra, Singpurwalla and Stephens (1983). The consideration of such problems is important in statistics primarily because of their appearance in the asymptotic theory of goodness-of-fit tests (see Chandra, Singpurwalla and Stephens, 1980 and 1981).

One of the highlighted illustrations from the 1983 Chandra, Singpurwalla and Stephens paper (henceforth called CSS) offers an

excellent example of the difficult search for a satisfactory experimental design and ultimately the need to make very long computer runs for adequate precision. This problem is the estimation by simulation of upper quantiles for the maximum of the (continuous-time) Brownian motion process over the unit interval (call it $W(t)$, $0 < t < 1$). We know that

$$E[W(t)] = 0 \qquad (\text{for all } t;\ W(0) = 0)$$

and that

$$E[W(t)W(s)] = \min(s,t) \qquad (0 \leq s,t \leq 1). \tag{1}$$

From a practice first used by Serfling and Wood, we can approximate this process at the points (j/k, k fixed and $0 \leq j \leq k$) by $(k+1)$-variate normal vectors with mean 0 and the same covariances as the Brownian motion. We know that

$$\lim_{k\to\infty} \Pr\{\max_{j\leq k} W(j/k) \geq x\} = 2\{1-\Phi(x)\} \qquad (x \geq 0), \tag{2}$$

where $\Phi(x)$ is the univariate normal CDF. Furthermore, Siegmund showed that the complementary CDF of the maximum could be well approximated for finite k as

$$2\{1-\Phi(x+.583/\sqrt{k})\} \quad , \tag{3}$$

with the result binding in the limit as $k \to \infty$. Note also that

the limiting density is the folded standard unit normal on $(0,\infty)$.

We often see estimation of both the population mean and variance from the same pseudo-random sample. However, the variance of a percentile is itself a function of the probability distribution of the variable. More completely, we know (for example, see Durbin, 1973) that the properly scaled and translated (np)th order statistic from an ordinary random sample has a limiting normal distribution, that is, that

$$\frac{X_{(np)}-\xi_p}{\sqrt{p(1-p)/n}} \xrightarrow{df} Z \sim N(0,1/f^2(\xi_p)) \quad (f = \text{original density}).$$

Since this variance of the upper quantiles depends on tail probabilities, it is clear that run size must rise rapidly as p goes to one for decaying densities. For example, the unit normal has approximate functional values of .1755 and .02665 for p=.9 and .99, respectively. Hence the estimator of the (.99n) order statistic has an approximate variance of 13.94/n as compared to 2.922/n for the (.9n) one, or a nearly five-fold increase. This further reinforces our earlier point of the need to make longer runs for the fixed precision estimation of tail quantiles.

Our target in this work then was to develop a more complete perspective on appropriate run sizes for this Gaussian problem. Of course, the exact critical values are known here; but this affords an excellent opportunity to calibrate any approach for determining sample sizes.

2. Results

CSS generated 10,000 (k+1)-variate normal vectors (k=20,30,50,60,90) using the extended precision version of the routine GGNSM from the International Mathematical and Statistical Library (IMSL). Their Table I shows the pth quantiles of the empirical distribution of

$$\max_{j \leq k} \{W(j/k)\}$$

from their simulation experiment for p =.9, .95, .975 and .99, using 10,000 replications. These results are compared to the corresponding quantiles given by Siegmund's approximation, and the authors claim to have encountered an important anomaly in such a simulation, namely, that there is a critical "turning down effect as k becomes large," which puts their point estimates on a path below what they should be. Even if we assume that the Siegmund approximation is close to correct for smaller k, there should be some variability about its values for any run size by the very nature of random draws. However, CSS felt that they had seen too much movement, and even a pattern down and away. But we observe that the kind of variation they noted is totally to be expected in light of what is really a rather small sample size, and the possible need for vastly different run sizes for different p values.

In our mind the problem of determining rates of convergence of sample quantiles for the maximum of a Brownian motion by simulation is very challenging and a good illustration for a more

general class of such problems. So we attempted to examine the problem more carefully and to develop a sounder strategy for finding an agreeable run size. We did this by repeating the basic CSS experimental design. We have created a FORTRAN77 program using GGNSM, with the same values of k, namely, 20, 30, 50, 60 and 90, and with p-values of .9, .95, .97 and .99. (We have opted to use .97 instead of .975 to keep the abscissa spacing uniform). Complete runs were performed for the eight sample sizes of 10,000, 20,000, 30,000, 40,000, 50,000, 70,000, 90,000 and 100,000.

Our simulation results are all documented in Tables I-VIII, where they are compared to those of CSS and what might be expected from the Siegmund approximation. To put all these numbers in clearer perspective, we have made two sets of plots. The first set is represented by Figures 1-3, where the quantiles for the three p values are plotted for the largest $k = 90$ and compared to their approximation given by (3). The second group provides a picture of the estimated rate of convergence of the critical values as k increases. As done by CSS, on each of these figures we have included a plot of the line which results for each quantile from Equation (3).

To summarize the salient insights offered by the first three figures, we note that the results are not at all surprising. For the moment assuming that the exact variabilities of the estimators are not known, each estimate (i.e., the .90, .95 and .99 quantiles) has become moderately stable with the increasing sample sizes (for $k = 90$). As expected, the .99 estimate is the least stable of the three. We might even find all of the 100,000

sample values acceptably precise to two or even three significant digits (but certainly no more than that).

However, this conclusion may be very misleading. A more completely constructed experimental design, arranged into blocks for estimating variances, would likely show a rather wide confidence interval for the .99 quantile, down to a fairly tight one for the .90th. In light of the fact that the underlying distribution theory is available for the Brownian maximum (that is, from Equation (2)), we can actually compute the variance of each of our three key quantiles. Since the density function of the maximum is the folded unit normal, we are easily able to compute the appropriate tail ordinates as

$$\begin{cases} .350996 & \text{for} \quad p = .90 \\ .206272 & \text{for} \quad p = .95 \\ .053304 & \text{for} \quad p = .99 \end{cases} .$$

Thus it follows that the standard deviations for the estimates of the .90, .95 and .99 quantiles taken from a sample of 100,000 are found as

$$\sqrt{\frac{p(1-p)}{nf^2(\xi_p)}} = \begin{cases} .0027028 & (p = .90) \\ .0033412 & (p = .95) \\ .0055903 & (p = .99) \end{cases} .$$

For purposes of understanding the precision of our estimators, let us overlay these (limiting) standard deviations onto the k = 90 final estimates of Figures 1-3 to make approximate (say, 95%) confidence statements. Thus the following intervals result (compared to Siegmund's estimate for k = 90):

(1.5893,1.6001) for p = .90 vs. Siegmund's value of 1.5835
(1.9026,1.9160) for p = .95 vs. Siegmund's value of 1.8985
(2.5174,2.5398) for p = .99 vs. Siegmund's value of 2.5385

Of course, we note the much wider interval for .99. In fact, we can calculate the sample size which would be needed to give the same absolute accuracy for the .99 estimate as in .90. This would be

$$N = \frac{(.99)(.01)}{(.0027028)^2(.053304)^2} \doteq 477{,}000,$$

or nearly <u>five</u> times the baseline sample.

For a sample of 10,000, the standard deviations for the quantile estimates increase $\sqrt{10}$-fold to

$$\sqrt{\frac{p(1-p)}{nf^2(\xi_p)}} = \begin{cases} .0085471 & (p=.90) \\ .0105659 & (p=.95) \\ .0176781 & (p=.99) \end{cases} .$$

The resultant confidence intervals (for any level) must thus be $\sqrt{10}$ times as wide. Thus, for example based on $k=90$ the actual .99 point would be estimated to be within (2.5428-.0353562, 2.5428+.0353562) or (2.5074, 2.5782), which is quite a broad interval. In our view, therefore, 10,000 is just too small a run size to give adequate precision, or indeed maybe to prevent any conclusion whatsoever.

The primary lessons of Figures 4-6 focus on the rates of convergence with respect to the parameter k. The work of

Siegmund clearly suggested that larger values of k should be used for more accurate estimation. But there is obviously a fairly large increase in computing time as the dimension of the multivariate normal goes up; but the payoff in precision could be significant, since simple linear extrapolation may not work well.

The result of Equation (3) translates into a straight line for each quantile when plotted against $1/\sqrt{k}=x$. For .90 the line is $y=1.645-.583x$, while for .99 it is $y=2.598-.583x$. Since the slopes are the same for all quantiles, each line reaches the y-axis at a point $.583/\sqrt{90} \doteq .0614$ higher than that for $x=1/\sqrt{90}$. Thus, for example, to reach a point within .01 of the limiting height would require a k of 58.3^2 or about 3,400. Furthermore, we notice that all 95% confidence intervals constructed for k=90 do not cover the actual answers. Hence, it seems clear that extreme caution should be used with this sort of approach.

To combine the results displayed in the figures, we might conclude that the .90 quantile is reasonably well approximated with a sample size of 100,000 and k = 90, though a somewhat larger k would be even better. For the .95 quantile, a slightly larger run size would be preferable, with k definitely increased beyond 90. And, finally, to get adequate precision for the .99 point, even larger increases would be warranted in the run size and k possibly up to as much as a sample of 500,000 and an order of magnitude (or higher) increase in k.

TABLE I

Estimated Percentage Points for Run Size = 10,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5367
	.95	1.9318	1.8296	1.8193
	.97	–	*	2.0275
	.99	2.5053	2.4696	2.4797
30	.90	1.6049	1.5386	1.5560
	.95	1.9362	1.8536	1.8571
	.97	–	*	2.0592
	.99	2.5815	2.4936	2.4536
50	.90	1.6376	1.5626	1.5938
	.95	1.9798	1.8776	1.8994
	.97	–	*	2.0697
	.99	2.5407	2.5176	2.4565
60	.90	1.5997	1.5697	1.5712
	.95	1.9416	1.8847	1.8767
	.97	–	*	2.0865
	.99	2.4642	2.5250	2.5264
90	.90	1.6239	1.5835	1.5893
	.95	1.8913	1.8985	1.8985
	.97	–	*	2.1189
	.99	2.4625	2.5385	2.5428

NOTE: The hyphens are in the .97 rows because CSS did not estimate for that p value, while we felt that such an estimate provided important insight. The asterisks are used to indicate that the approximation can be used for .97, but we chose not to do the calculation since a full comparison is not possible.

TABLE II

Estimated Percentage Points for Run Size = 20,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5352
	.95	1.9318	1.8296	1.8429
	.97	–	*	2.0525
	.99	2.5053	2.4696	2.4700
30	.90	1.6049	1.5386	1.5271
	.95	1.9362	1.8536	1.8402
	.97	–	*	2.0760
	.99	2.5815	2.4936	2.4918
50	.90	1.6376	1.5626	1.5594
	.95	1.9798	1.8776	1.8612
	.97	–	*	2.1042
	.99	2.5407	2.5176	2.5044
60	.90	1.5997	1.5697	1.5533
	.95	1.9416	1.8847	1.8599
	.97	–	*	2.0954
	.99	2.4642	2.5250	2.5102
90	.90	1.6239	1.5835	1.5838
	.95	1.8913	1.8985	1.8777
	.97	–	2.1885	2.0691
	.99	2.4625	2.5385	2.4787

TABLE III

Estimated Percentage Points for Run Size = 30,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5312
	.95	1.9318	1.8296	1.8419
	.97	-	*	2.0541
	.99	2.5053	2.4696	2.4679
30	.90	1.6049	1.5386	1.5361
	.95	1.9362	1.8536	1.8584
	.97	-	*	2.0505
	.99	2.5815	2.4936	2.4630
50	.90	1.6376	1.5626	1.5613
	.95	1.9798	1.8776	1.8731
	.97	-	*	2.0742
	.99	2.5407	2.5176	2.4971
60	.90	1.5997	1.5697	1.5690
	.95	1.9416	1.8847	1.8923
	.97	-	*	2.0932
	.99	2.4642	2.5250	2.4837
90	.90	1.6239	1.5835	1.5666
	.95	1.8913	1.8985	1.8892
	.97	-	2.1885	2.0927
	.99	2.4625	2.5385	2.5068

TABLE IV

Estimated Percentage Points for Run Size = 40,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5281
	.95	1.9318	1.8296	1.8356
	.97	-	*	2.0505
	.99	2.5053	2.4696	2.4679
30	.90	1.6049	1.5386	1.5375
	.95	1.9362	1.8536	1.8404
	.97	-	*	2.0514
	.99	2.5815	2.4936	2.5039
50	.90	1.6376	1.5626	1.5594
	.95	1.9798	1.8776	1.8707
	.97	-	*	2.0739
	.99	2.5407	2.5176	2.4918
60	.90	1.5997	1.5697	1.5666
	.95	1.9416	1.8847	1.8768
	.97	-	*	2.0837
	.99	2.4642	2.5250	2.5144
90	.90	1.6239	1.5835	1.5846
	.95	1.8913	1.8985	1.8977
	.97	-	2.1885	2.1053
	.99	2.4625	2.5385	2.5082

TABLE V

Estimated Percentage Points for Run Size = 50,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5284
	.95	1.9318	1.8296	1.8358
	.97	–	*	2.0505
	.99	2.5053	2.4696	2.4648
30	.90	1.6049	1.5386	1.5304
	.95	1.9362	1.8536	1.8441
	.97	–	*	2.0660
	.99	2.5815	2.4936	2.4595
50	.90	1.6376	1.5626	1.5667
	.95	1.9798	1.8776	1.8781
	.97	–	*	2.0801
	.99	2.5407	2.5176	2.4777
60	.90	1.5997	1.5697	1.5638
	.95	1.9416	1.8847	1.8832
	.97	–	*	2.0982
	.99	2.4642	2.5250	2.4852
90	.90	1.6239	1.5835	1.5797
	.95	1.8913	1.8985	1.8894
	.97	–	*	2.0933
	.99	2.4625	2.5385	2.4919

TABLE VI

Estimated Percentage Points for Run Size = 70,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5264
	.95	1.9318	1.8296	1.8338
	.97	-	*	2.0505
	.99	2.5053	2.4696	2.4547
30	.90	1.6049	1.5386	1.5410
	.95	1.9362	1.8536	1.8501
	.97	-	*	2.0608
	.99	2.5815	2.4936	2.4731
50	.90	1.6376	1.5626	1.5620
	.95	1.9798	1.8776	1.8857
	.97	-	*	2.0895
	.99	2.5407	2.5176	2.4772
60	.90	1.5997	1.5697	1.5704
	.95	1.9416	1.8847	1.8878
	.97	-	*	2.0921
	.99	2.4642	2.5250	2.4996
90	.90	1.6239	1.5835	1.5972
	.95	1.8913	1.8985	1.9099
	.97	-	*	2.1177
	.99	2.4625	2.5385	2.5283

TABLE VII

Estimated Percentage Points for Run Size = 90,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5257
	.95	1.9318	1.8296	1.8324
	.97	-	*	2.0505
	.99	2.5053	2.4696	2.4582
30	.90	1.6049	1.5386	1.5417
	.95	1.9362	1.8536	1.8562
	.97	-	*	2.0656
	.99	2.5815	2.4936	2.4732
50	.90	1.6376	1.5626	1.5688
	.95	1.9798	1.8776	1.8856
	.97	-	*	2.0966
	.99	2.5407	2.5176	2.4940
60	.90	1.5997	1.5697	1.5695
	.95	1.9416	1.8847	1.8909
	.97	-	*	2.1056
	.99	2.4642	2.5250	2.5153
90	.90	1.6239	1.5835	1.5960
	.95	1.8913	1.8985	1.9127
	.97	-	*	2.1312
	.99	2.4625	2.5385	2.5344

TABLE VIII

Estimated Percentage Points for Run Size = 100,000

Values of k	Values of p	CSS Estimate	From Approx	Harris Estimate
20	.90	1.5725	1.5146	1.5240
	.95	1.9318	1.8296	1.8320
	.97	-	*	2.0492
	.99	2.5053	2.4696	2.4565
30	.90	1.6049	1.5386	1.5437
	.95	1.9362	1.8536	1.8613
	.97	-	*	2.0695
	.99	2.5815	2.4936	2.4691
50	.90	1.6376	1.5626	1.5646
	.95	1.9798	1.8776	1.8796
	.97	-	*	2.0906
	.99	2.5407	2.5176	2.4873
60	.90	1.5997	1.5697	1.5774
	.95	1.9416	1.8847	1.9037
	.97	-	*	2.1120
	.99	2.4642	2.5250	2.5192
90	.90	1.6239	1.5835	1.5947
	.95	1.8913	1.8985	1.9093
	.97	-	*	2.1193
	.99	2.4625	2.5385	2.5286

Figure 1

.90 QUANTILE FOR MAX OF BROWNIAN MOTION (K=90)

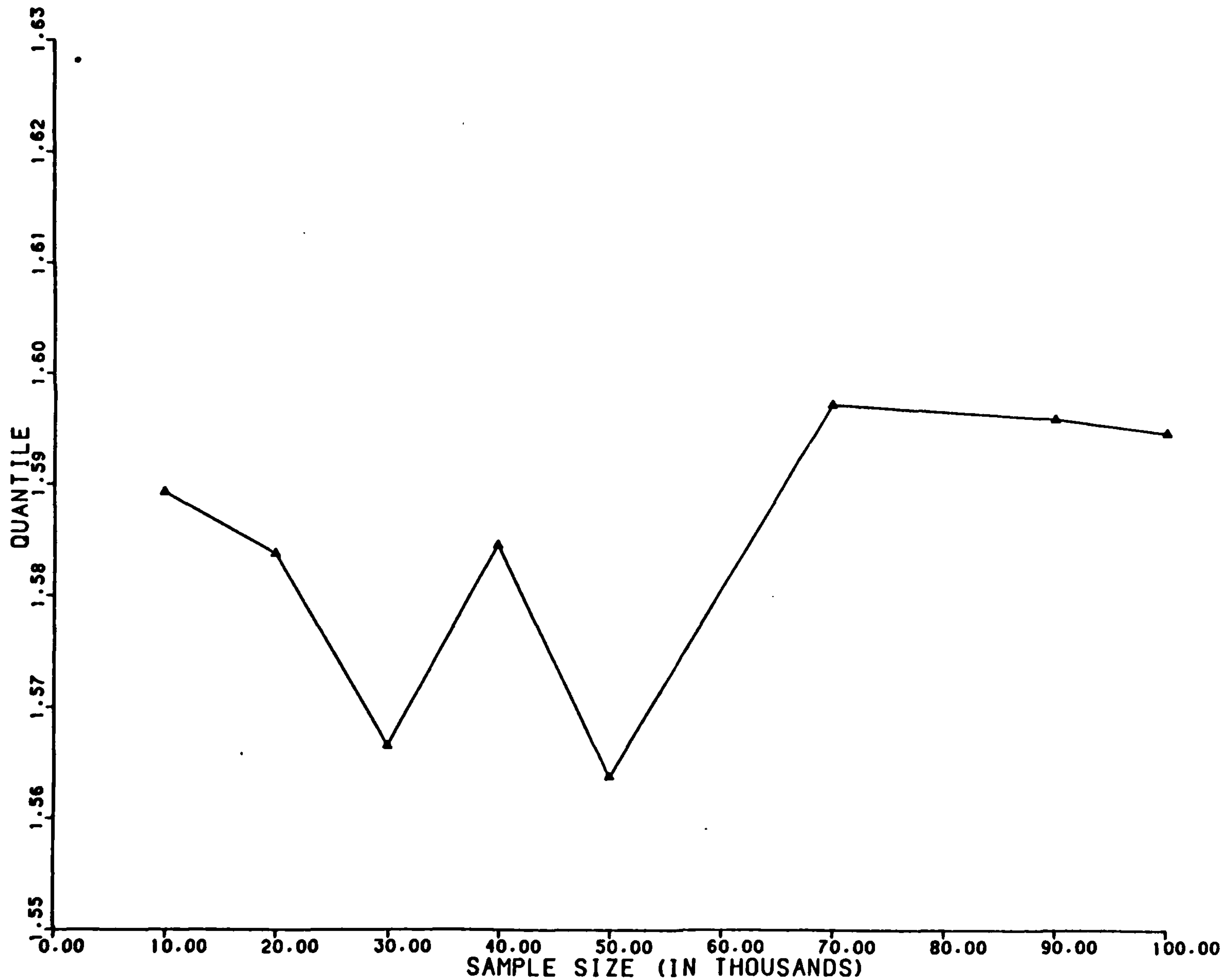


Figure 2

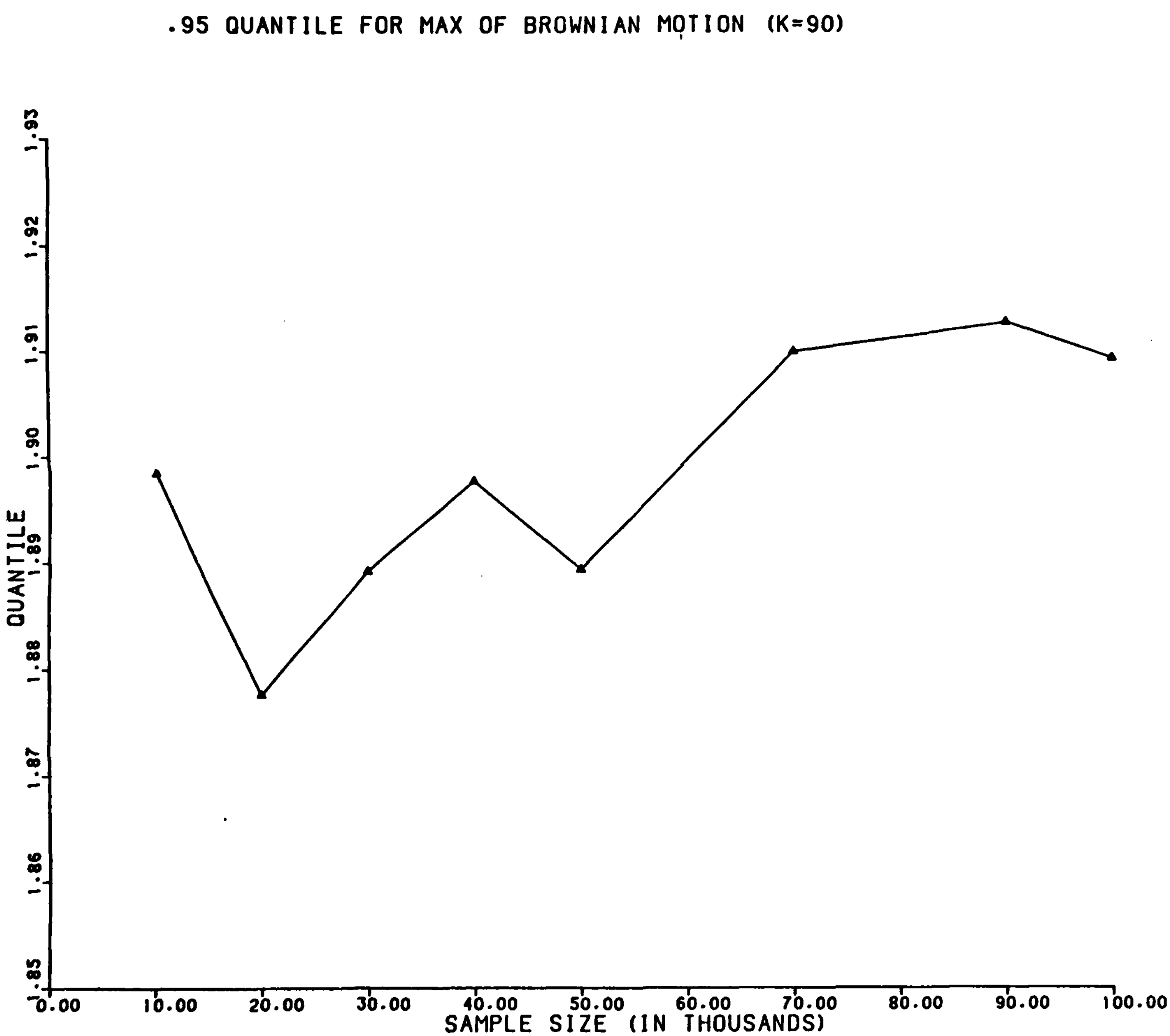

.95 QUANTILE FOR MAX OF BROWNIAN MOTION (K=90)
QUANTILE
1.85
1.86
1.87
1.88
1.89
1.90
1.91
1.92
1.93
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
100.00
SAMPLE SIZE (IN THOUSANDS)

Figure 3

.99 QUANTILE FOR MAX OF BROWNIAN MOTION (K=90)

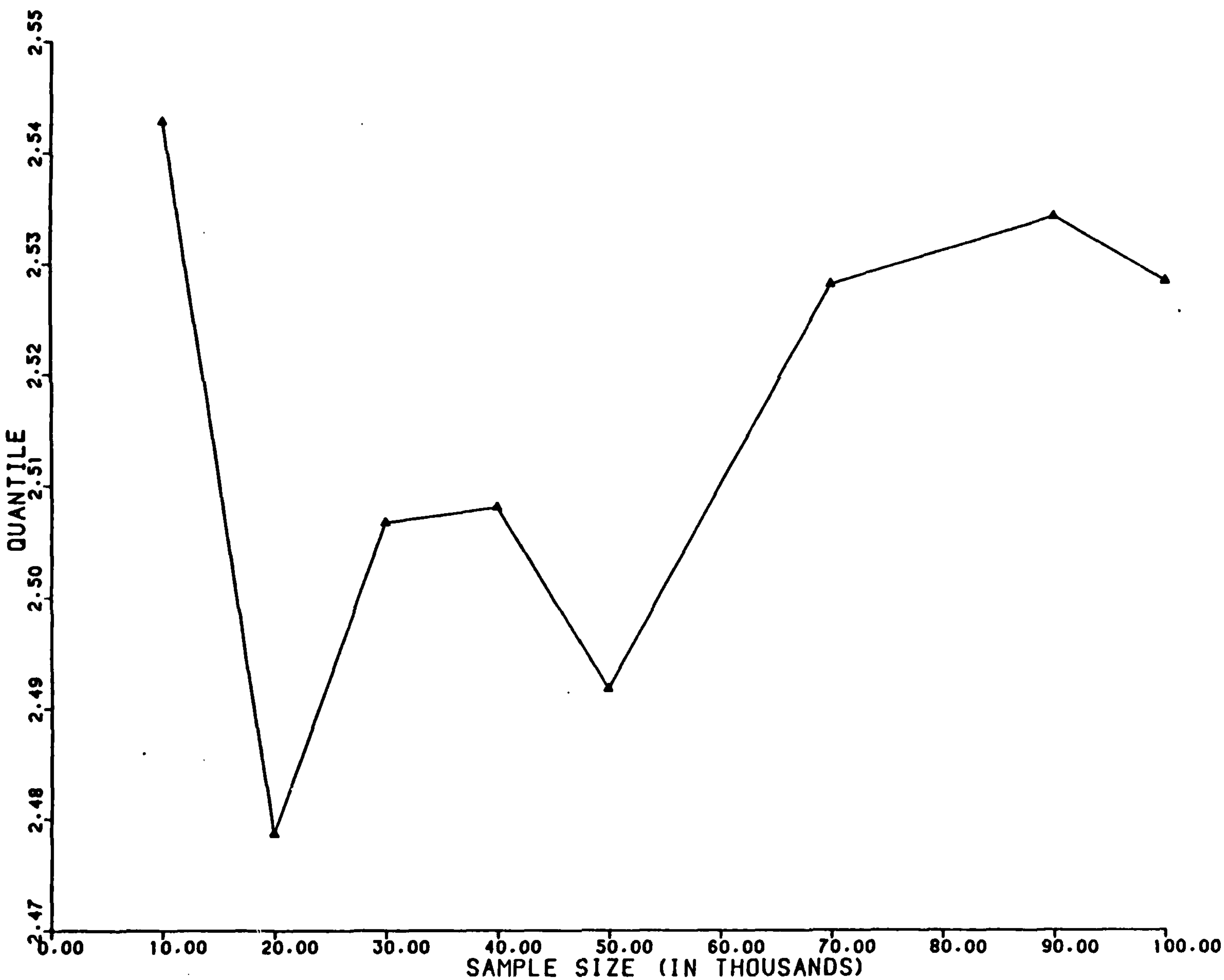


Figure 4

.90 QUANTILES FOR SAMPLE SIZES OF 100.000

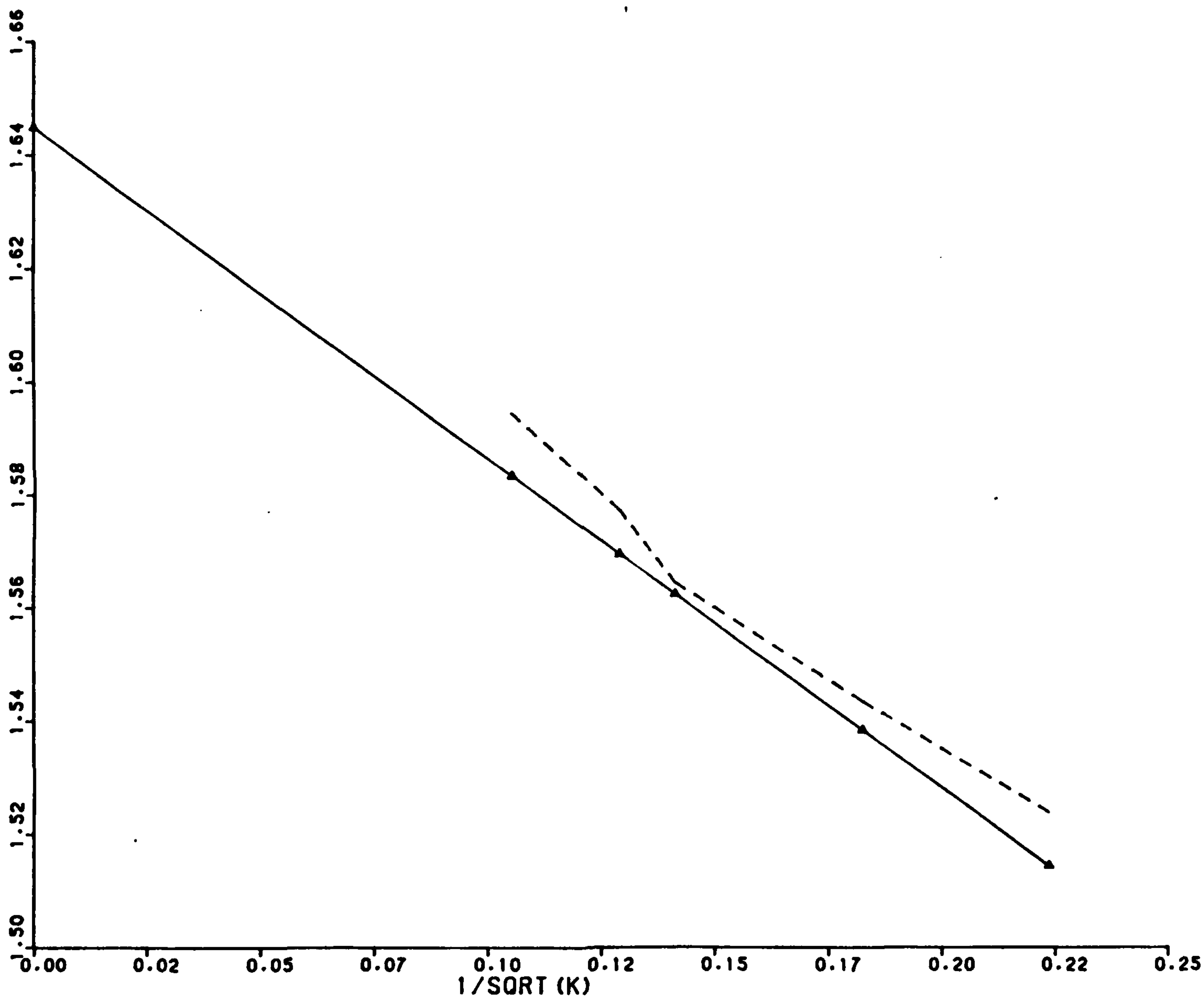


Figure 5

.95 QUANTILES FOR SAMPLE SIZES OF 100.000

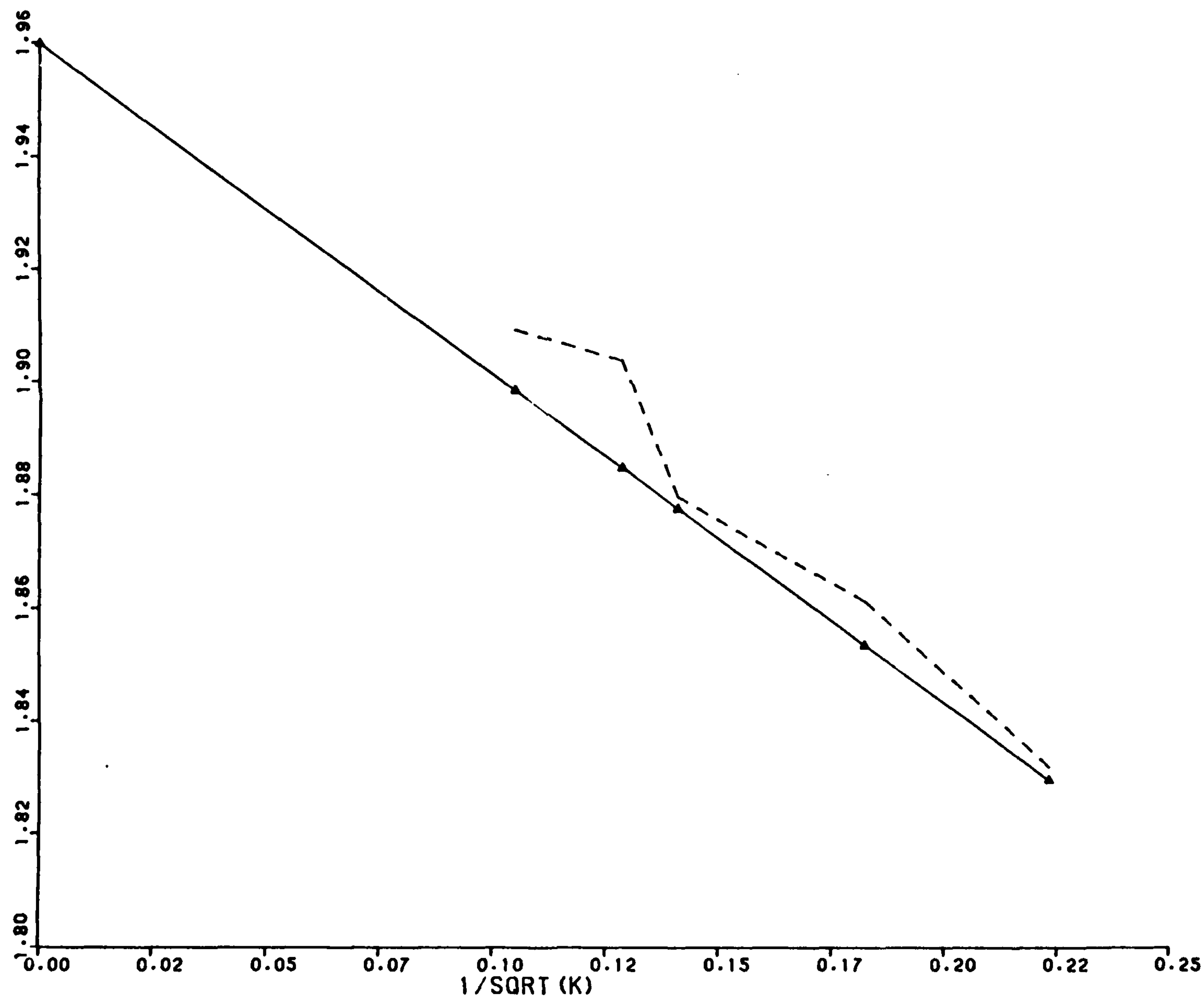
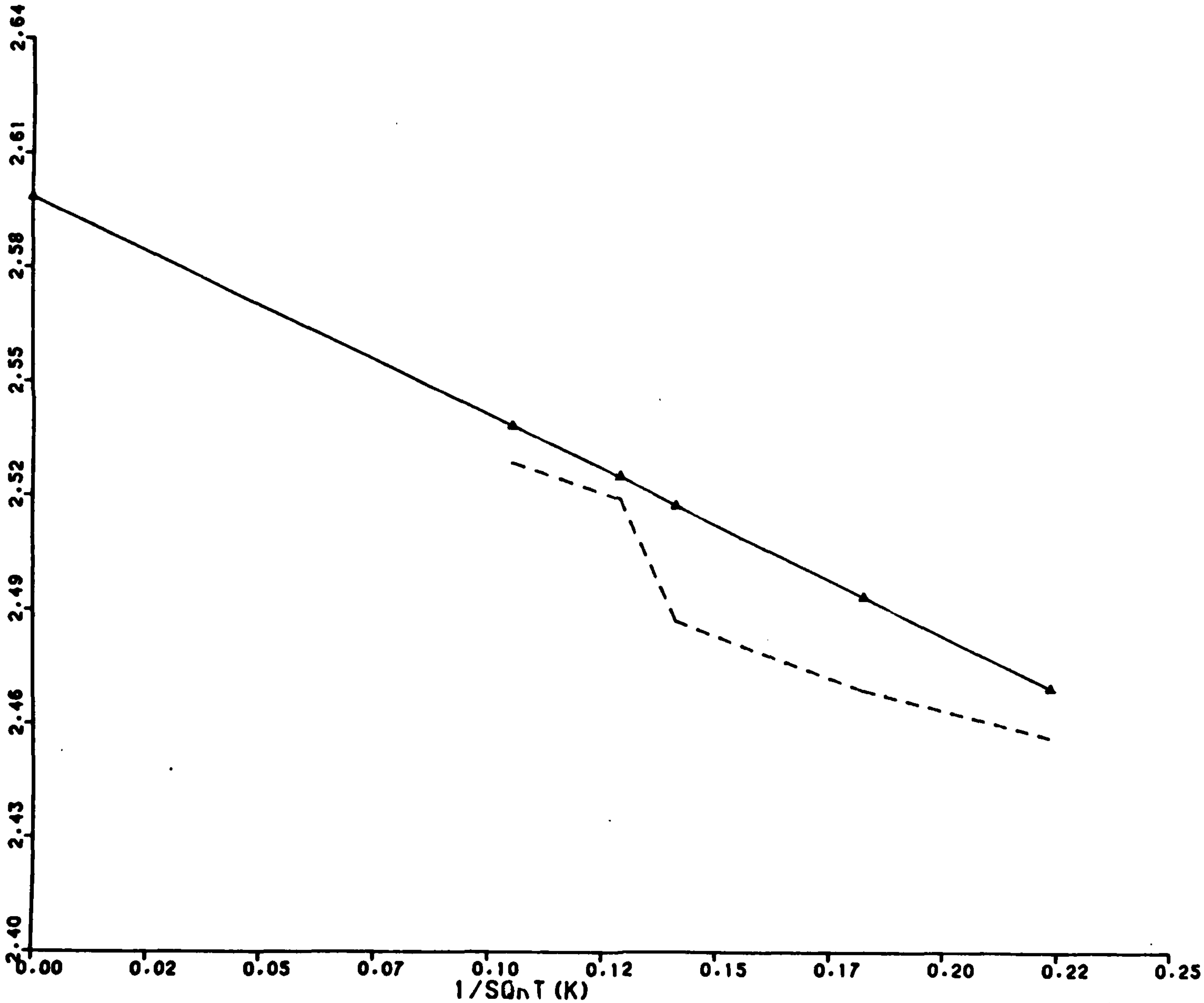


Figure 6

.99 QUANTILES FOR SAMPLE SIZES OF 100,000



Acknowledgements

The author wishes to express his sincere thanks to Camille A. Harris (no relation) for her able help in preparing the software for this analysis. Her work was performed as a fourth-year thesis submitted to the faculty of the University of Virginia in partial fulfillment of the requirements for the B.S. degree in Computer Science. Further thanks go to Nozer D. Singpurwalla for his interest in having more work done on this problem.

REFERENCES

Bratley, P., B.L. Fox and L.E. Schrage (1983), A Guide to Simulation, Springer-Verlag, New York.

Chandra, M., N.D. Singpurwalla and M.A. Stephens (1980), Goodness-of-fit tests for the Weibull and the extreme value distributions with estimated parameters, Technical Report, Department of Statistics, Stanford University.

Chandra, M., N.D. Singpurwalla and M.A. Stephens (1981), Kolmogorov statistics for tests of fit for the extreme value and Weibull distributions, J. Amer. Statist. Assoc., 76, 375, 729-731.

Chandra, M., N.D. Singpurwalla and M.A. Stephens (1983), Some problems in simulating the quantiles of the maxima and other functionals of Gaussian processes, J. Statist. Comput. Simul., 18, 45-57.

Durbin, J. (1973), Weak convergence of the sample distribution function when parameters are estimated, Ann. Statist., 1, 279-290.

Rubinstein, R.Y. (1981), Simulation and the Monte Carlo Method, John Wiley and Sons, Inc., New York.

Siegmund, D. (1978), Corrected diffusion approximation in certain random walk problems, Technical Report No. 4, Department of Statistics, Stanford University, Stanford, CA.

Serfling, R.J. and C.L. Wood (1976), On the null hypothesis limiting distributions of Kolmogorov-Smirnov type statistics with estimated location and scale parameters, Technical Report, Florida State University, Tallahassee, FL.

Wood, C.L. (1978), A large sample Kolmogorov-Smirnov test for normality of experiment error in a randomized block design, Biometrika, 65, 673-676.

DISTRIBUTION LIST

Copy No.	
1 - 20	Office of Naval Research 800 North Quincy Street Arlington, VA 22217 Attention: Program Manager Statistics and Probability
21	ONR Resident Representative Joseph Henry Building Room 623 2100 Pennsylvania Avenue, N.W. Washington, D.C. 20037
22*	Defense Technical Information Center Building 5, Cameron Station Alexandria, VA 22314
23 - 24	C. M. Harris
25	C. C. White
26 - 27	E. H. Pancake Sci./Tech. Information Center
28	SEAS Publications Files

*Reproducible copy

JO# 5937:wn

UNIVERSITY OF VIRGINIA
School of Engineering and Applied Science

The University of Virginia's School of Engineering and Applied Science has an undergraduate enrollment of approximately 1,500 students with a graduate enrollment of approximately 500. There are 125 faculty members, a majority of whom conduct research in addition to teaching.

Research is a vital part of the educational program and interests parallel academic specialties. These range from the classical engineering disciplines of Chemical, Civil, Electrical, and Mechanical and Aerospace to newer, more specialized fields of Biomedical Engineering, Systems Engineering, Materials Science, Nuclear Engineering and Engineering Physics, Applied Mathematics and Computer Science. Within these disciplines there are well equipped laboratories for conducting highly specialized research. All departments offer the doctorate; Biomedical and Materials Science grant only graduate degrees. In addition, courses in the humanities are offered within the School.

The University of Virginia (which includes approximately 1,500 full-time faculty and a total full-time student enrollment of about 16,000), also offers professional degrees under the schools of Architecture, Law, Medicine, Nursing, Commerce, Business Administration, and Education. In addition, the College of Arts and Sciences houses departments of Mathematics, Physics, Chemistry and others relevant to the engineering research program. The School of Engineering and Applied Science is an integral part of this University community which provides opportunities for interdisciplinary work in pursuit of the basic goals of education, research, and public service.

END

FILMED

6-85

DTIC